GPS LOCATIONS FOR GIS: GETTING THEM RIGHT THE FIRST TIME

Caroline Erickson and Pierre Héroux
Geodetic Survey Division, Geomatics Canada
Natural Resources Canada
615 Booth Street
Ottawa, Ontario
K1A 0E9
caroline@geod.emr.ca
heroux@geod.emr.ca

ABSTRACT

The Global Positioning System (GPS) is a very powerful tool which has, and will continue to provide the location data for attributes in many Geographical Information Systems (GIS). Manufacturers of GPS equipment have developed low cost, user-friendly receivers and software which are capable of easily satisfying many users' positioning requirements. Assuming the successful operation of GPS receivers and software, other measures are necessary to ensure the integrity of positions being entered into a GIS. This paper addresses one very essential consideration in using GPS for GIS: coordinate systems. As well, three positioning techniques suitable for many GIS are discussed and compared: single point positioning; differential positioning; and single point positioning with precise orbits and clocks. Results based on this last, recently introduced technique show that metre level accuracies may be achieved with a single receiver. Finally, a few suggestions are offered for novice users to help get GPS positioning right the first time.

INTRODUCTION

Location data is a fundamental element of Geographical Information Systems (GIS). More and more, GPS is being used as the tool for GIS georeferencing. The successful application of GPS is dependent on understanding its coordinate system, achievable accuracies and limitations.

This paper explains the significance of the GPS coordinate system and how large positional errors in a GIS may be avoided through careful attention to coordinate systems.

All GPS positioning is not the same. Accuracies ranging from 100 m to millimetres may be achieved depending on the data collection and processing techniques, and the hardware and software used. As can be expected, in general, the higher the accuracy, the higher the GPS positioning cost. This paper describes three lower accuracy techniques, based on code (also called pseudorange) measurements, which are suitable for GIS applications.

Prepared for Decision 2001, Toronto Ont., September 12-16, 1994.

One of these techniques, single point positioning with precise ephemerides and clocks, is particularly noteworthy as it is new to the GPS positioning community (Héroux et al., 1993) and has significant benefits for many GIS applications. An awareness of the different GPS positioning alternatives is important so that the technique most appropriate for a specific requirement is applied.
The paper concludes with comments on the limitations of GPS positioning and suggestions to help ensure the successful application of GPS for GIS georeferencing.

COORDINATE SYSTEMS

There are currently 25 GPS satellites orbiting 20,000 km above the earth, configured in a constellation such that at least four satellites are visible anywhere in the world at anytime. These satellites broadcast data which is collected by GPS receivers and used for computing positions. In a simplified sense, positioning with GPS is achieved by simultaneously recording the signals from at least four satellites at the user's receiver. By knowing the locations of the satellites in the sky (referred to as orbits or ephemerides), the point of intersection of the four signals may be computed. It is therefore necessary to relate the coordinates of the GPS satellites to the coordinates used daily in GIS (Figure 1).

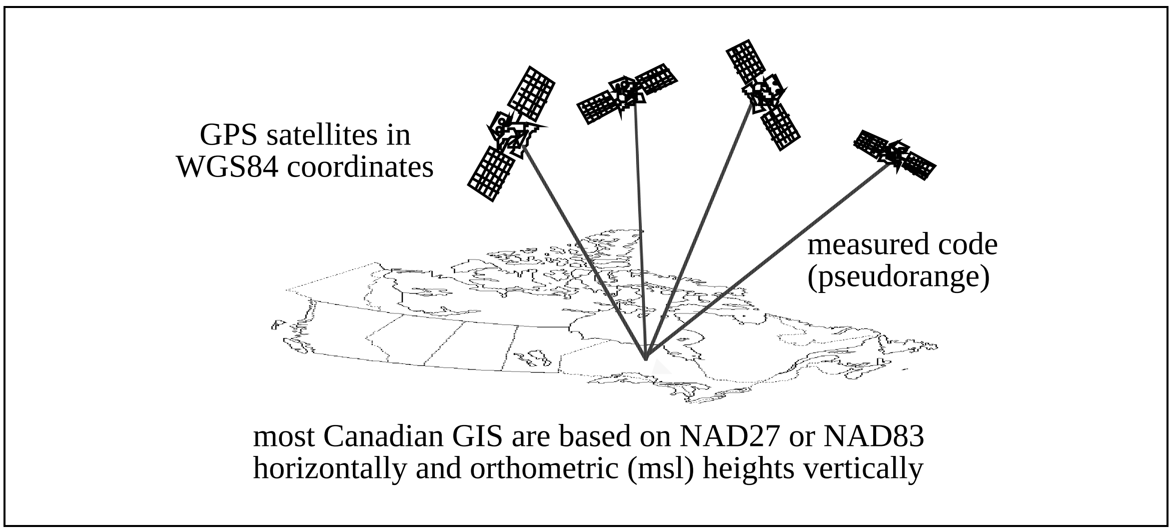

Figure 1. Single Point Positioning and Coordinate Systems

The orbits of the GPS satellites are referenced to WGS84 (World Geodetic System 1984) and are usually represented in Cartesian coordinates (x,y,z) such that the origin is at the centre of the earth, the x axis passes through the Greenwich Meridian and the z axis passes through the north pole.

In GIS, horizontal coordinates are usually represented as latitudes and longitudes, or UTM northings and eastings, based on a reference datum such as NAD27 or NAD83. Each of these datums is in turn based on an ellipsoid which represents the size and shape of the earth. An ellipsoid is a smooth mathematical surface which can be thought of as a sphere that is "squashed" at the poles. Points on the ellipsoid may be represented as Cartesian coordinates or as latitudes, longitudes and ellipsoidal heights (i.e. height above the ellipsoid).

In GIS, vertical coordinates are usually represented as heights above mean sea level, which are referred to as orthometric heights.

Horizontal Coordinates

Horizontal coordinates based on the North American Datum of 1983 (NAD83) are fully compatible with the WGS84 GPS based coordinates. The same cannot be said for the forerunner to NAD83, the North American Datum of 1927 (NAD27). Shifts in geodetic coordinates resulting from the transition from NAD27 to NAD83 range from about 120 metres westerly on the west coast to 70 metres easterly in Newfoundland and 100 metres northerly in the high Arctic as shown in Figure 2. The corresponding Universal Transverse Mercator (UTM) coordinates have a fairly consistent northward shift ranging from about 200 to 250 metres (Pinch, 1990). The shifts are a result of using an earth-centred ellipsoid consistent with WGS84 instead of a non-geocentric ellipsoid as used in NAD27, and removal of distortions in the NAD27 coordinates through a complete readjustment.

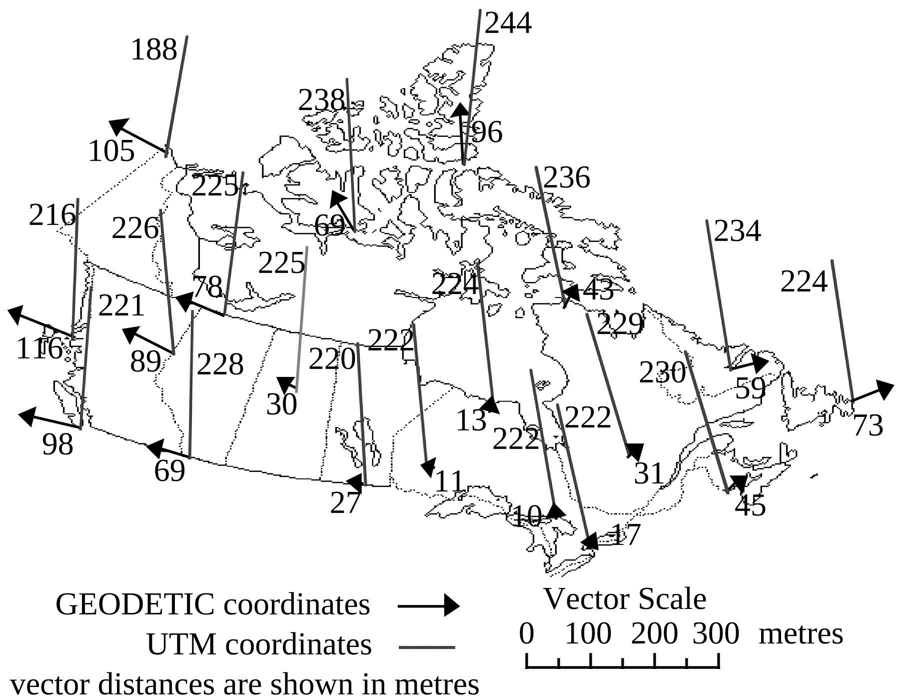

Figure 2. NAD27 - NAD83 Coordinate Difference Vectors (adapted from Pinch, 1990)

Users of GPS must ensure data collected in WGS84 coordinates is used in an NAD83 based GIS, or transformed to be compatible with the GIS datum. Failing to take this step could introduce very large errors into a data base.

Note that some GPS receiver manufacturers include software to carry out these transformations. Usually, such software packages only take into account the shifts between the two reference ellipsoids, and not the removal of NAD27 distortions. Consequently errors up to 20 m may still remain in some transformation packages (Junkins, 1990). The user must ascertain if such transformations are sufficiently accurate for their needs. Transformation software between NAD27 and NAD83 which more fully models the actual local shifts is available from provincial survey agencies and the Geodetic Survey Division of Geomatics Canada (see Appendix).

Vertical Coordinates

The link between the ellipsoidal heights derived from GPS and the mean sea level (orthometric) heights, which we commonly deal with, is more complex than the datum relationship for horizontal coordinates.

Orthometric heights actually are referenced to the geoid. The geoid is defined as an equipotential surface (i.e. a surface on which the gravity potential is constant) that closely represents mean sea level. It forms a smooth but irregular surface around the earth that differs significantly from the geometrically defined ellipsoid.

The ellipsoidal and orthometric heights are linked by the geoid height as illustrated in Figure 3. The geoid height may be obtained from a geoid model related to the WGS84 ellipsoid such as that computed by the Geodetic Survey Division, referred to as GSD91 (Véronneau and Mainville, 1992). It is therefore possible to obtain orthometric heights from ellipsoidal heights with knowledge of the geoid height at a location, as interpolated from a geoid model. (Note that some GPS receiver manufacturers provide geoid models from which orthometric heights are computed. The user must ascertain if the embedded geoid models are sufficiently accurate for their needs.)

ellipsoidal height	(from GPS)
- geoid height	(from geoid model)
= orthometric height	

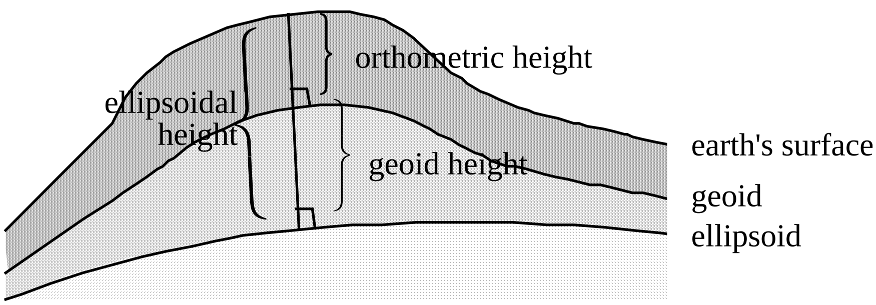

Figure 3. How to Derive Orthometric Heights from GPS

GPS users operating in single point positioning mode for height determination must ascertain whether ellipsoidal or orthometric heights are produced by their receiver's processing software. Geoid heights (i.e. the difference between ellipsoidal and orthometric heights) in Canada vary up to ±50 m, depending on location, which is therefore the maximum error one would expect if ellipsoidal heights were mistakenly taken as orthometric heights.

Note that for relative positioning where the position of one point is determined with respect to another, an orthometric height may be derived with smaller errors if the orthometric height of the reference point is known. This is because the relative differences between the ellipsoid and geoid are much smaller than the absolute differences between the two surfaces.

For GPS positioning, it is important to be aware of coordinate systems (i.e. NAD83 coordinates or others, orthometric or ellipsoidal heights) and to take the appropriate steps to ensure the right data is entered into a GIS. (Further information on coordinate systems, the geoid and their significance for GPS positioning may be found in Erickson, 1993.)

Most professional surveyors are familiar with these concepts and should be consulted by GIS developers who are exploring the use of GPS.

GPS POSITIONING METHODS FOR GIS

GPS positioning accuracies ranging from centimetres to 100 m may be desired, depending on the specific GIS requirements. This paper will limit discussion to lower accuracies (1 to 100 m) based on using code measurements, which are often more economically viable for GIS applications. Single point positioning, differential positioning and single point positioning with precise orbits and clocks will be discussed.

Single Point Positioning

Single point positioning is achieved by intersecting the measurements from four or more satellites at a single receiver on the earth's surface. The accuracies achievable using single point positioning are 100 m 2drms horizontally and 156 m 2σ vertically assuming favourable satellite geometry. These accuracies apply equally to static or kinematic single point positioning. Solutions may be attained almost instantaneously, using an inexpensive GPS receiver.

The limited accuracy achievable with single point positioning is mainly due to inaccuracies in the broadcast satellite orbits and clocks. In addition, the delay of the signals as they travel through the earth's ionosphere and troposphere reduces accuracy. Multipath (i.e. the reception of signals which reflect off ground surface objects rather than traveling directly to the antenna) and receiver noise (i.e. the receiver's limitations in accurately measuring the code) also affects resulting accuracies.

The U.S. Department of Defense, which maintains the GPS satellites, intentionally degrades the broadcast satellite orbits and dithers the satellite clocks as part of an official policy to limit positioning accuracy for unauthorized users. This degradation is referred to as Selective Availability (S/A).

Two techniques may be used to improve the cited 100 m accuracies to the 1 to 10 m level: (1) differential positioning, and (2) single point positioning with precise orbits and clocks. In each method, improved results are achieved by greatly reducing some of the above mentioned error sources, including the effects of S/A.

Differential GPS Positioning

Differential positioning may be conducted with either post-mission or real-time kinematic processing. The former is simpler and less expensive, while the latter is complicated by the requirement for a real-time data link. Using differential positioning, the coordinates of one point which is used as a base station must be known. The difference between the measured satellite ranges at this base station and the "true" satellite ranges is computed to produce a correction, which is then applied to measured ranges at a second point. The corrected ranges are used to compute coordinates in a single point positioning algorithm. The effect of orbit and satellite clock errors, as well as atmospheric delays, are greatly reduced using this technique, leading to accuracies ranging from 1 to 10 m.

Some of the factors which affect whether 1 m or 10 m accuracy is achievable using differential GPS include the distance from the base receiver, receiver characteristics (noise, multipath resistance), and satellite geometry.

Differential GPS may be carried out by using two receivers with one serving as a base station, using one receiver and subscribing to a service that supplies differential GPS corrections (Reason, 1994) or using one receiver with data collected at a Canadian Active Control System site (see Figure 4). Differential GPS is a well-accepted technique which has been applied in the GPS community for several years.

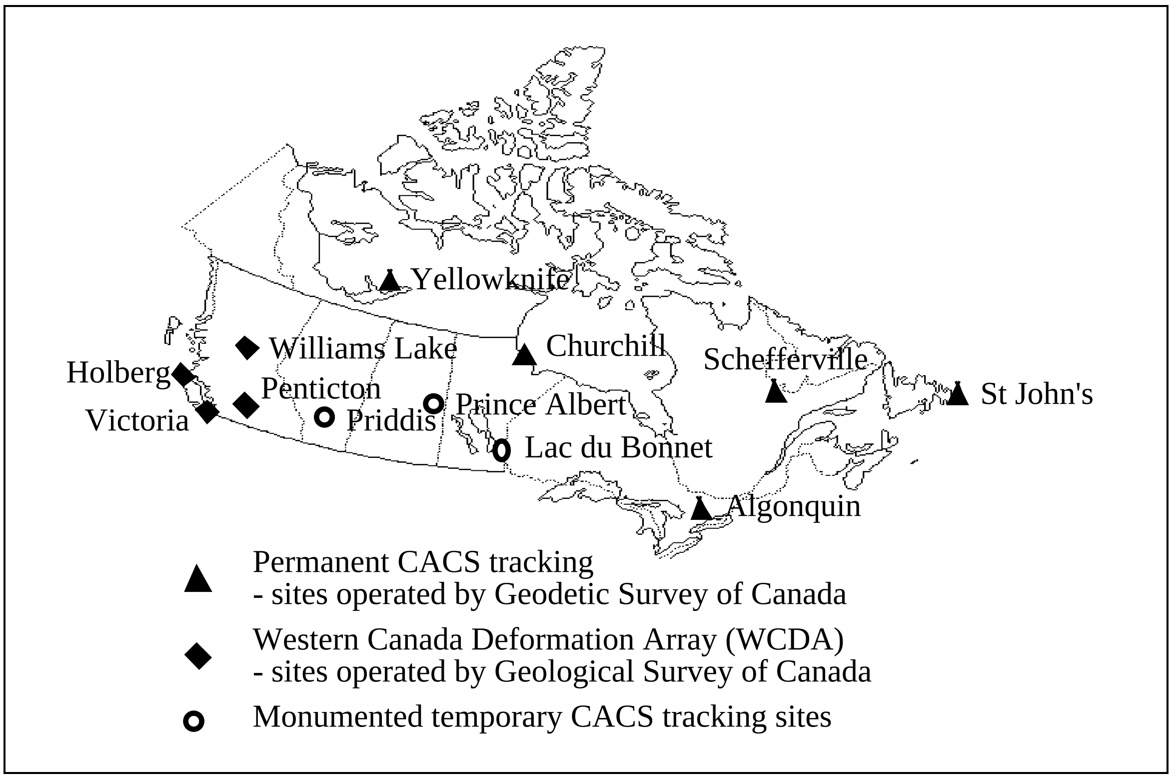

Figure 4. Canadian Active Control System Network Configuration

A variation of differential GPS which may be encountered, particularly in navigation applications, is wide area differential GPS (WADGPS). In this technique, information to produce range corrections is derived from multiple base stations instead of just a single base station (Mueller, 1994). With WADGPS the requirement of being close to a base station to achieve higher accuracies is greatly reduced.

Single Point Positioning with Precise Orbits and Clocks

Single point positioning with precise orbits and clocks is similar to differential positioning in that large orbital and clock errors are greatly reduced, however the approach used is significantly different. Instead of determining range corrections at a base station and applying them at the point requiring coordinates, the broadcast orbit and clock parameters are replaced by precise orbits and clock parameters, which are available as a product of the Canadian Active Control System (CACS). The differences between single point positioning, differential positioning and single point positioning with CACS precise orbits and clocks are illustrated in Figure 5.

Single Point Positioning	broadcast orbits & clocks measured pseudoranges	~ 100 m accuracy
Differential Positioning	broadcast orbits & clocks measured pseudoranges + corrections from base station	~ 1 to 10 m accuracy
Single Point Positioning with CACS Precise Orbits and Clocks	CACS precise orbits & clocks measured pseudoranges	~ 1 to 10 m accuracy

Figure 5. Code Positioning Techniques

A description of the Canadian Active Control System is necessary to understand the source of the precise orbits and clocks.

The Canadian Active Control system consists of unattended GPS tracking stations distributed across the country (see Figure 4) referred to as Active Control Points (ACPs). These ACPs continuously record carrier phase and pseudorange measurements for all

GPS satellites within station view. The data collected is retrieved on a daily basis by a central processing facility in Ottawa. The main objectives of the CACS are (1) to provide direct and convenient access to the Canadian Spatial Reference System and (2) to improve effectiveness and precision of GPS applications. This is accomplished by providing a fiducial reference frame, by computing and making available precise satellite ephemerides (orbital parameters) and precise satellite clock corrections; by monitoring GPS integrity and performance from the analysis of data acquired through continuous tracking and; by supporting Differential GPS development and other services (geodynamics, precise time transfer, etc.).

The CACS orbits, which are also based on a number of global sites, have an accuracy of approximately 10 cm, which is far superior to the 5 - 10 m accuracy of the GPS broadcast orbits. Similarly the CACS precise clocks are accurate to 1 nanosecond (30 cm), whereas the broadcast clocks are only accurate to 70 - 100 nanoseconds (21 - 30 m).

Single point positioning with precise orbits and clocks was recently introduced in the GPS community (Heroux et al., 1993; Lachapelle et al., 1994). Program *GPSPACE* (GPS Positioning from ACS Clocks and Ephemerides) was developed by the Geodetic Survey of Canada as an interface program for CACS products (see Appendix). The program performs standard single point positioning (i.e. with broadcast orbits and clocks), single point positioning with precise orbits and clocks, as well as differential positioning with respect to a single base station. As single point positioning with precise orbits and clocks is new to the GPS community, it is worthwhile to show some results demonstrating its capabilities compared to use of broadcast orbits.

Figures 6 and 7 show the quality of positioning possible at a single epoch with a single receiver when using either the broadcast orbits and clocks or the CACS precise orbits and clocks. Program *GPSPACE* was used in post-processing. The data set was collected in static mode at a 1 second rate at station "Babbage" in Ottawa on August 4, 1994 between 12:49:55 and 13:15:14 UT using a NovAtel GPS receiver.

In Figure 6, the difference in latitude and longitude from the known coordinates is shown for single point positioning solutions using a) broadcast orbits and clocks, and b) CACS precise orbits and clocks. The plots show independent one second solutions for each orbit type. Errors up to 60 metres in latitude and 25 metres in longitude are evident when broadcast orbits and clocks are used, compared to errors of less than three metres in each direction when CACS precise orbits and clocks are used.

In Figure 7, the difference from the known ellipsoidal height is shown. The height discrepancies, which range up to 140 m when broadcast orbits and clocks are used, are limited to a few metres through use of CACS precise orbits and clocks.

Comparison of Code Positioning Techniques

Of the three code positioning techniques described, each has advantages and disadvantages as listed in Table 1. It is up to users, to be knowledgeable about the different options, and choose which technique is most appropriate for their application.

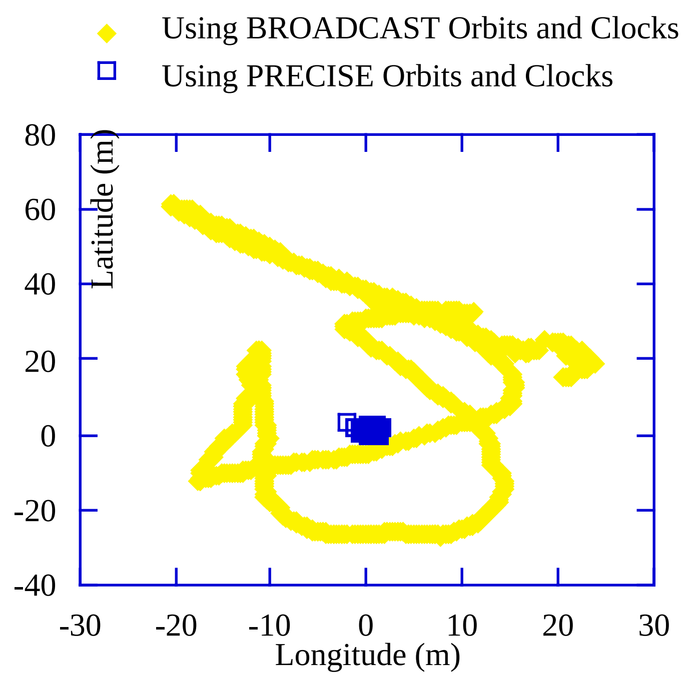

Figure 6. Comparison of Latitude and Longitude of Single Point Solutions using a) broadcast orbits and clocks and b) CACS precise orbits and clocks.

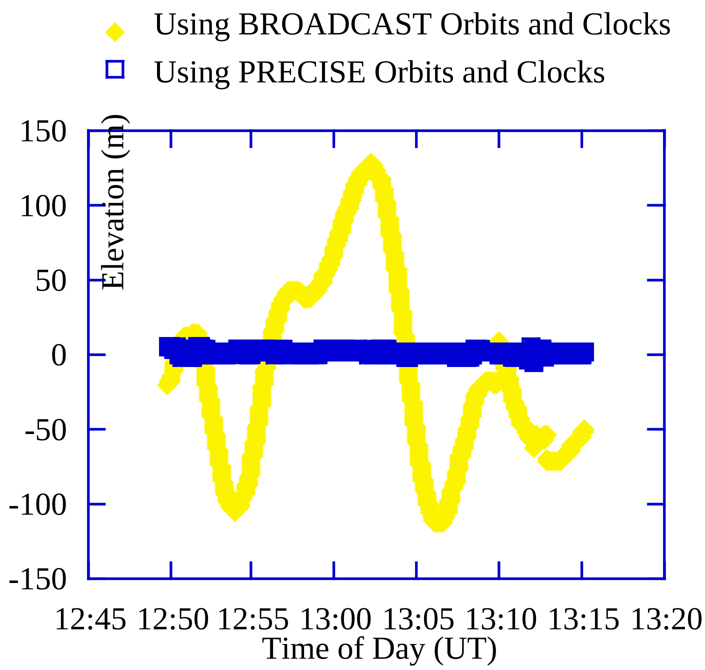

Figure 7. Comparison of Heights of Single Point Solutions using a) broadcast orbits and clocks and b) CACS precise orbits and clocks.

Table 1. Advantages and Disadvantages of Three Code Positioning Techniques

Single Point Positioning (with broadcast orbits and clocks)	*advantages* • base station not required • static or kinematic • real time or post-mission • uses all satellites in view *disadvantages* • low accuracy (~ 100 m)
Differential Positioning (based on code measurements)	*advantages* • static or kinematic • real time or post-mission • higher accuracy (1 to 10 m) *disadvantages* • base station required • must match correction and measurement files • can only use satellites which are also in view at base station • data link required for real-time • accuracy dependent on distance from base station
Single Point Positioning with CACS Precise Orbits and Clocks	*advantages* • base station not required • static or kinematic • higher accuracy (1 to 10 m) • automatically tied to Canadian Spatial Reference System • applicable for all of Canada • uses all satellites in view *disadvantages* • currently only post-mission processing possible

GPS FOR GIS - GETTING IT RIGHT

Up to this point, the importance of the horizontal and vertical coordinate systems and a discussion of three code positioning techniques suitable for GIS applications have been presented. A few other items are worth mentioning to help get GPS for GIS right the first time.

Users should always be aware of one of the greatest limitations of GPS, that is the requirement that the satellite-receiver path be unobstructed. It is impossible to use GPS in tunnels or underpasses, very difficult amongst urban high-rises, and difficult in forested areas. It is for this reason that several positioning systems offer integration of GPS with complementary technologies.

With GIS, attribute data is typically collected and tagged with a position. For optimal efficiency and accuracy, time tagging of the measured attribute and the position data should be automatically linked (Gilbert, 1994).

The most important step in ensuring success of GPS for GIS is to test the full system from data collection to final processing. The equipment used for GPS varies greatly in complexity, cost and capabilities. It cannot be assumed that all accuracy claims given by equipment manufacturers or other users will be consistently met under all production field conditions, and hence it is important to test and evaluate the equipment. For the same reasons it is also important to test and evaluate GPS processing software and techniques. The position integrity should be tested by using points of known superior accuracy, and the fit with the GIS application should be tested with field trials.

GPS is a tremendously powerful, efficient and effective positioning tool, which when used with care, will form a major input for a GIS system.

ACKNOWLEDGEMENTS

The insights and guidance provided by Jan Kouba in the concept and development of program *GPSPACE* is gratefully acknowledged. Thanks are extended to Pierre Sauvé who collected the GPS data which was presented in this paper. Don Junkins is acknowledged for his input on the National Transformation. A special recognition goes to members of the Canadian Active Control System Team, who ensure the continual successful operation of the system and the timely production of precise orbits and clocks.

REFERENCES

Erickson, C. 1993. GPS Positioning Guide. Geodetic Survey Division, Natural Resources Canada.

Fenton, P., B. Falkenberg, T. Ford, K. Ng and A.J. Van Dierendonck 1991. NovAtel's GPS Receiver - The High Performance OEM Sensor of the Future. Proceedings of ION GPS'91, Albuquerque, NM., Sept. 10-13, The Institute of Navigation, Washington, D.C., pp. 49-58.

Gilbert, C. 1994. Integrating Other Measuring Devices With GPS Positions. Earth Observation Magazine. June 1994. pp. 51-53.

Héroux, P., M. Caissy, and J. Gallace (1993). Canadian Active Control System Data Acquisition and Validation. Proceedings of the 1993 IGS (International GPS Service for Geodynamics) Workshop, University of Bern, pp. 49-58.

Junkins, D. 1990. The National Transformation for converting between NAD27 and NAD83 in Canada. Moving to NAD '83 the new address for georeferenced data in Canada. The Canadian Institute of Surveying and Mapping. Ottawa, Ont., pp. 16-40.

Lachapelle, G., Klukas, R., Qiu, W. and Melgard, T.E. 1994. Single Point Satellite Navigation Accuracy - What the Future May Bring. Presented at IEEE PLANS94, Las Vegas, 11-15 April 1994.

Pinch, M.C. 1990. Differences Between NAD27 and NAD83. Moving to NAD '83 the new address for georeferenced data in Canada. The Canadian Institute of Surveying and Mapping. Ottawa, Ont., pp. 1-15.

Reason, T. 1994. Differential for a Dollar a Day. Professional Surveyor. July/August, pp. 4-10.

Mueller, T. 1994. Wide Area Differential GPS. GPS World, June, pp. 36-44.

Véronneau, M. and A. Mainville 1992. Computation of a Canadian Geoid Model Using FFT Technique to Evaluate Stokes' and Vening-Meinesz' Formulas in a Planar Approximation, Internal Report, Geodetic Survey Division, Canada.

APPENDIX

Requests for information concerning

(1) the Canadian Active Control System and its products (including program *GPSPACE*)
(2) the Canadian Geoid Model GSD91
(3) federal horizontal and vertical control
(4) The National Transformation for converting between NAD27 and NAD83 in Canada

should be addressed to:

Canada Centre for Surveying
Geodetic Survey Division
Information Services
615 Booth Street
Ottawa, Ontario K1A 0E9
(613) 995-4410 or (613) 992-2061
FAX: (613) 995-3215
E.Mail: information@geod.emr.ca